The design and characterization of two proteins with 88% sequence identity but different structure and function

Patrick A. Alexander, Yanan He, Yihong Chen, John Orban, and Philip N. Bryan*

Center for Advanced Research in Biotechnology, University of Maryland Biotechnology Institute, 9600 Gudelsky Drive, Rockville, MD 20850

Edited by David Baker, University of Washington, Seattle, WA, and approved June 6, 2007 (received for review January 31, 2007)

To identify a simplified code for conformational switching, we have redesigned two natural proteins to have 88% sequence identity but different tertiary structures: a 3-α helix fold and an α/β fold. We describe the design of these homologous heteromorphic proteins, their structural properties as determined by NMR, their conformational stabilities, and their affinities for their respective ligands: IgG and serum albumin. Each of these proteins is completely folded at 25°C, is monomeric, and retains the native binding activity. The complete binding epitope for both ligands is encoded within each of the proteins. The IgG-binding epitope is functional only in the α/β fold, and the albumin-binding epitope is functional only in the 3-α fold. These results demonstrate that two monomeric folds and two different functions can be encoded with only 12% of the amino acids in a protein (7 of 56). The fact that 49 aa in these proteins are compatible with both folds shows that the essential information determining a fold can be highly concentrated in a few amino acids and that a very limited subset of interactions in the protein can tip the balance from one monomer fold to another. This delicate balance helps explain why protein structure prediction is so challenging. Furthermore, because a few mutations can result in both new conformation and new function, the evolution of new folds driven by natural selection for alternative functions may be much more probable than previously recognized.

evolution | folding | NMR | protein design | protein structure

How amino acid sequence determines protein structure and ultimately protein function is perhaps the most fundamental unresolved question in biology. The degenerate nature of folding information, the vast number of conformations available to a polypeptide chain, and the low stability of most natural proteins ($\Delta G_{unfolding}$ between 5 and 15 kcal/mol) all present challenges to understanding the folding code. In this study we create a set of proteins in which the essential folding information for two alternative folds resides within a small number of amino acids.

The starting point in this process was protein G, the multidomain, cell wall protein from *Streptococcus* (Lancefield group G) (1). Protein G contains two types of domains that bind to serum proteins in blood: G_A domains of 47 structured amino acids that bind to HSA (2, 3) and G_B domains of 56 structured amino acids that bind to the constant (Fc) region of IgG (4, 5). The ability to bind serum proteins apparently confers selective advantage to pathogenic bacteria by allowing them to camouflage themselves with host proteins (6).

The G_A and G_B domains share no significant sequence homology and have different folds, 3-α and α/β, respectively. We have previously characterized 56-aa versions of the G_A and G_B domains. High-resolution structures of these proteins have been determined, and the energetics of their folding and ligand binding reactions have been measured by microcalorimetry and hydrogen–deuterium exchange. The G_A domain used in this study (PSD1) binds to HSA with a $K_d = 20$ nM and has a $\Delta G_{unfolding}$ of 6 kcal/mol (25°C, 0.1 M KPO_4, pH 7.2) (7). Amino acids 1–7 and 55–56 are disordered in G_A (PSD1). The remaining 47 aa are well ordered in a 3-α helix bundle (8) (Fig. 1*A*). The G_B domain used in this study (G_{B1}) binds to the constant (Fc) region of IgG with a $K_d = 100$ nM and has a $\Delta G_{unfolding}$ of 7 kcal/mol (25°C, 0.1 M KPO_4, pH 7.2) (9, 10). All 56 aa of G_{B1} are well ordered in a four-stranded β-sheet with an α-helix connecting strands two and three (11) (Fig. 1*B*). Both solution and crystal structures have been determined for G_B (11–15).

The Fold-Specific Folding Problem

G_A and G_B constitute a heteromorphic pair of proteins (two proteins of equal length that have two different folds). We wish to study the structural and energetic relationship between the two distinct folded states for this pair. Consider the equilibrium shown in Scheme 1.

For the native G_A sequence, only the horizontal limb of the equilibrium is detectable because the α/β fold is poorly populated. An identical equilibrium exists for the native G_B sequence, but only the vertical limb of the equilibrium is populated. The 47 positions of nonidentity in this initial heteromorphic pair shift the equilibrium from 99.996% 3-α fold (6 kcal/mol) and undetectable α/β to 99.999% α/β (7 kcal/mol) and undetectable 3-α. It is well known that not every amino acid in a protein contributes equally toward specifying the native fold, however (16). Thus, not all of the nonidentities will have equal information content toward specifying one fold or the other.

To identify a simplified code for conformational switching, we have determined mutations in binary sequence space (choice of either the G_A or G_B amino acid) which can be introduced into each member of the heteromorphic pair while preserving its native fold. Thus, new heteromorphic pairs of increasing identity are generated. The positions of nonidentity in the heteromorphic pair of highest identity constitute an essential fold-specific folding code parsed from the overall stability code. This basic approach has been taken in a number of previous studies (17–21). We have been able to build on the previous work to create heteromorphic pairs of higher identity and higher stability while preserving biological function.

Results

Encoding Latent Binding Sites. The first step in the process of creating homologous versions of G_A and G_B was mutating key amino acids in each so that the IgG- and HSA-binding epitopes are encoded in both proteins. The IgG-binding epitope is functional in the α/β fold and latent in the 3-α fold, whereas the albumin-binding epitope is functional in the 3-α fold and latent in the α/β fold (Fig.

Author contributions: P.A.A., J.O., and P.N.B. designed research; P.A.A., Y.H., Y.C., and J.O. performed research; P.A.A., Y.H., J.O., and P.N.B. analyzed data; and P.N.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

*To whom correspondence should be addressed. E-mail: bryan@umbi.umd.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0700922104/DC1.

© 2007 by The National Academy of Sciences of the USA

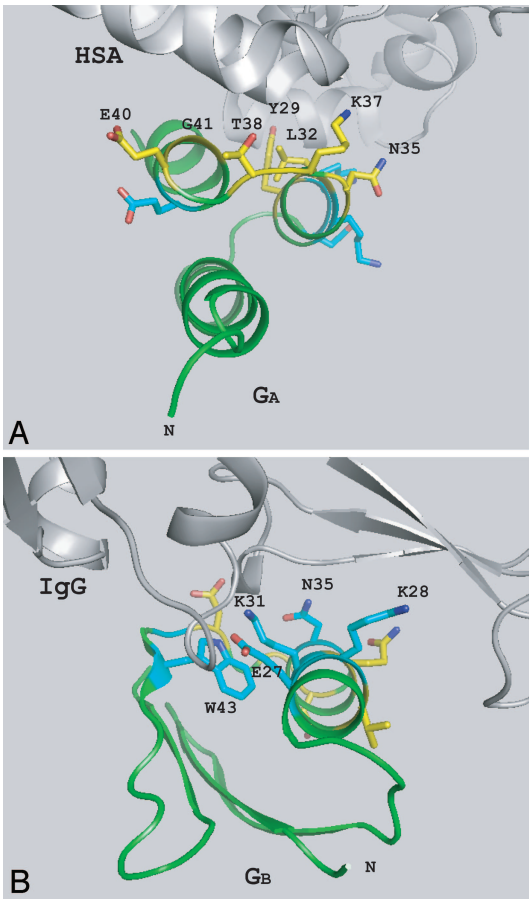
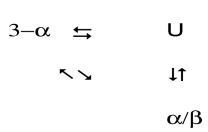


Fig. 1. Binding epitopes of G_A and G_B. (*A*) HSA-binding epitope of G_A. Shown is a cartoon depiction of the secondary structure of G_A. Amino acids making direct contacts with HSA (gray) are shown in yellow. Amino acids that were mutated to create a latent IgG epitope are in cyan (from Protein Data Bank entry 1TFO) (22). (*B*) IgG(Fc)-binding epitope of G_B. Shown is a cartoon depiction of the secondary structure of G_B. Amino acids making direct contacts with Fc (gray) are shown in cyan. Amino acids that were mutated to create a latent HSA epitope are in yellow (from Protein Data Bank entry 1fcc) (40).

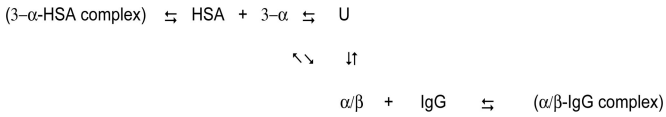
1). This was done for two reasons. First we wished to determine whether a latent binding function could be encoded without altering native activity. This finding, in itself, demonstrates that latent functions could be encoded in alternative folds of other proteins. Second, we wished to create an experimental system in which binding function can be used to detect a specific fold in future studies. Consider the equilibrium below in which binding function is thermodynamically linked to a particular conformational state (see Scheme 2).

If both binding epitopes are present in both members of a heteromorphic pair, we can potentially detect a poorly populated fold through its thermodynamic linkage to the bound form.

The HSA binding site in G_A comprises amino acids Y29, L32, N35, K37, T38, E40, and G41 (7, 22) (Fig. 1*A*). The IgG binding site in G_B comprises amino acids E27, K28, K31, and N35 on the central helix, W43 in the β3-strand, and main chain contacts in the intervening turn (3, 23, 24) (Fig. 1*B*). A latent IgG-binding epitope



Scheme 1.



Scheme 2.

was encoded in G_A with three mutations: D27E, Y28K, and E43W. K31 and N35 are already present in G_A. The resulting mutant is denoted G_A30.

A latent HSA binding site was encoded in G_B with the mutations V29Y, Q32L, N37K, G38T, and D40E. N35 and G41 are already present in G_B. The resulting mutant is denoted G_B30. G_A30 and G_B30 are 30% identical (Fig. 2).

G_A30 and G_B30 were characterized by CD to determine secondary structure content, thermal denaturation to determine conformation stability, gel filtration to confirm that the proteins remain monomeric, and affinity chromatography on IgG and HSA Sepharose to assess binding affinities. The CD spectra of G_A30 and G_B30 are shown in Fig. 3*A*. The CD spectrum of G_A30 is essentially identical to its parent (PSD1), and the CD spectrum of G_B30 is essentially identical to that of native G_{B1}. Thermal denaturation profiles followed by ellipticity at 222 nm are shown in Fig. 3*B*. The thermal denaturation midpoint is 86°C for G_A30 and 65°C for G_B30. The CD melting data were converted to $\Delta G_{\text{unfolding}}$ by using the relationship $\Delta G_{\text{unfolding}} = -\text{RTln}([\text{unfolded}]/[\text{folded}])$.

A plot of ΔG vs. temperature for G_A30 and G_B30 is shown in Fig. 4. G_A30 is similar in stability to its parent (PSD1). Fig. 4*A* shows the stability vs. temperature graph for G_A30 determined from the CD data. For reference, the figure also shows the stability profile of PSD1 determined from differential scanning calorimetry and hydrogen–deuterium exchange experiments (7, 8). The introduction of the latent IgG binding site into PSD1 does not have any significant affect on the thermodynamics of the unfolding reaction, including any variation in the change in heat capacity for unfolding (ΔCp). ΔCp is correlated with the change in solvent-exposed surface area upon unfolding and determines the curvature of the stability curve. The absence of any change in ΔCp is an indication that no major perturbations in packing of the hydrophobic core have occurred (25).

Fig. 4*B* shows the stability vs. temperature graph for G_B30 determined from the CD data. For reference, the figure also shows the stability profile of G_{B1} determined from differential scanning calorimetry and hydrogen–deuterium exchange experiments (9, 10). G_B30 is less stable than G_{B1} by almost 3 kcal/mol but nevertheless exhibits stability typical of some natural IgG-binding domains. Thus, latent function can be built into both proteins in a way that is relatively benign in terms of stability.

The binding of G_A30 and G_B30 to IgG and HSA was determined by using HSA and IgG Sepharose columns. The concentrations of the two ligands on the respective columns were similar to their concentrations in human serum (10 mg/ml). The use of the affinity columns allowed rapid assessment of binding affinity (and no binding interaction), as well as the stoichiometry of binding (two G_B molecules per IgG and one G_A molecule per HSA). Retention on the column after washing with 20-column volumes indicates a dissociation constant of <1 μM. Elution from a column with no tailing of the elution peak indicates a dissociation constant of >1 mM. Approximately 1 mg of G_A30 was tightly bound to HSA Sepharose and required 0.5 M NaOAc (pH 3.0) for elution. G_A30 eluted as a sharp peak in the flow-through fractions from IgG Sepharose. In contrast, $\approx$0.7 mg of G_B30 was tightly bound to IgG Sepharose and required 0.5 M NaOAc (pH 3.0) for elution. G_B30 eluted as a sharp peak in the flow-through fractions from HSA Sepharose. Thus, the engineering of latent binding function into the two proteins was accomplished without significant alteration of the native binding function. As expected, the IgG-binding epitope is

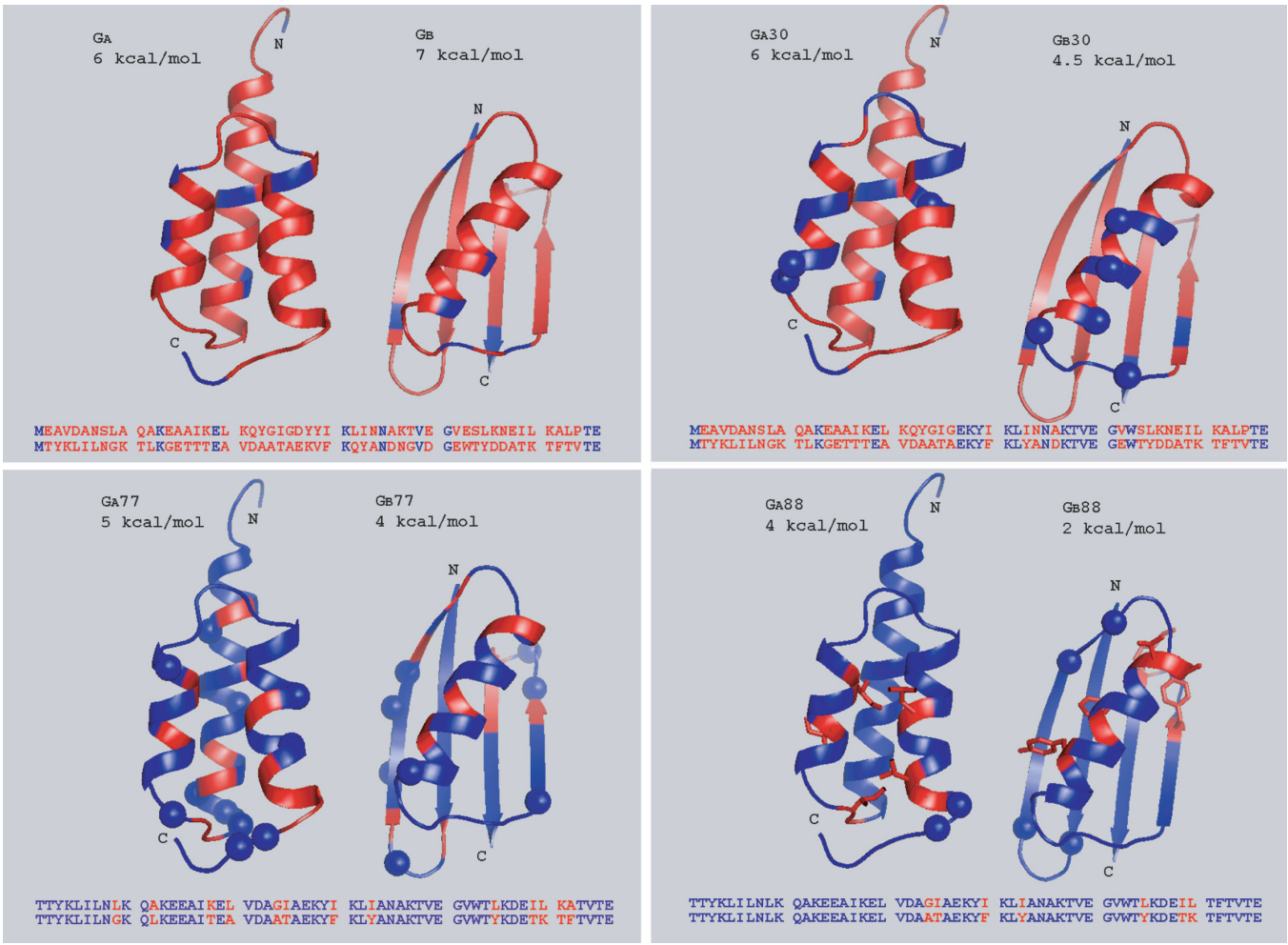


Fig. 2. Iterative design of heteromorphic pairs. For each heteromorphic pair, amino acid identities are shown in blue and nonidentities are shown in red. Mutations introduced in each design cycle are shown as blue spheres in the corresponding panel. The seven unique amino acids in G_A88 and G_B88 are shown as red sticks in *Lower Right*. Stabilities were measured in 0.1 M potassium phosphate buffer (pH 7.2) and extrapolated to 25°C (Fig. 4). The cartoons were generated from Protein Data Bank entries 2FS1 (PSD1) and 1PGA (G_{B1}).

cryptic in G_A30 because its correct presentation requires the α/β fold, and the HSA-binding epitope is cryptic in G_B30 because its correct presentation requires the 3-α fold.

Finally, gel filtration experiments on G25 and G75 columns were used to confirm that G_A30 and G_B30 had the same hydrodynamic properties as the parents, PSD1 and G_{B1}.

Increasing Identity. G_A30 and G_B30 constitute a heteromorphic pair with wild-type binding function, as well as excess functional capacity in latent form. We needed to methodically examine the binary sequence space at the 39 positions of nonidentity to determine whether highly homologous regions of that sequence space encode for both of the folds.

Previous studies with G_B have shown that this is not trivial. The tolerance of any particular mutation in a fold is context-dependent and cannot be determined by making sequential single amino acid substitutions. Blanco *et al.* (20) performed a methodical analysis of sequence space between G_B and the all-β SH3 domain of spectrin. They minimized the context-dependence issue by avoiding mutations of residues in the hydrophobic cores. Several mutants of high identity were found in their sampling that had circular dichroic spectra indicative of alternative folds and cooperative thermal denaturation transitions typical of folded proteins. None of the homologous pairs was found to have a well defined tertiary structure as judged by ^{1}H NMR, however. Dalal and colleagues (21, 26) carefully designed hybrid sequences of G_B and the all α-helical ROP homodimer. They were able to create heteromorphic pairs of up to 80% identity. The ROP-like proteins were disordered at neutral pH but acquired α-helical CD spectra below pH 5 and exhibit cooperative thermal unfolding transitions. Their limited solubility precluded detailed analysis of their tertiary structures (26). We have previously used phage display selection to explore the sequence space between G_B and the α-helical IgG-binding domain of protein A. A heteromorphic pair of 59% identity was evolved (27). High-resolution NMR structures of this pair are described by He *et al.* (28). Above 60% identity, the folding signal degraded to the point where the stability of heteromorphic pairs dropped below 2 kcal/mol and detailed structural analysis became problematic.

To find mutation-tolerant sites in G_A30 and G_B30 we relied heavily on previous experiments that used random mutagenesis of topological islands and phage display to select for functional folds (7, 8, 27). These experiments allowed us to classify each position in G_A and each position in G_B into one of three general categories: (*i*) mutation-tolerant and context-independent (mutations are tolerated independent of neighboring amino acids); (*ii*) mutation-tolerant but context-dependent (mutations are tolerated but only with compensating mutation of neighboring amino acids); and (*iii*) mutation-resistant (mutations seldom or never appear in selected

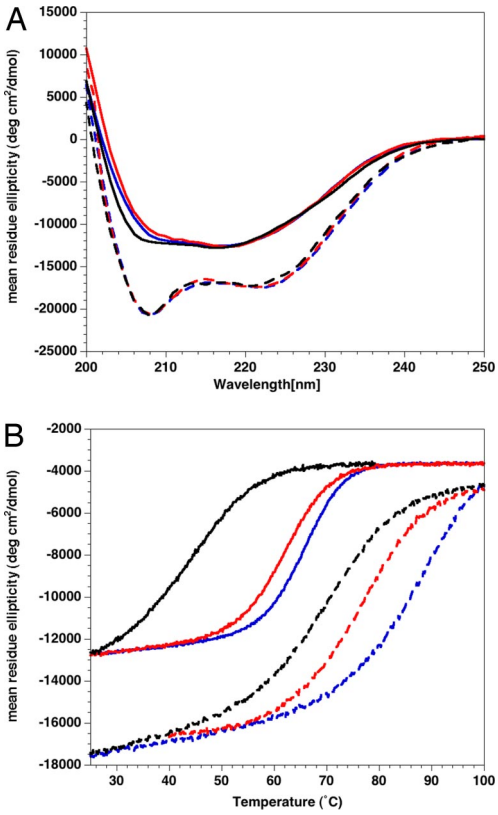


Fig. 3. Analysis of conformation and thermal denaturation by CD. CD spectra and thermal denaturation curves are shown for G_A30 (blue dashed line), G_A77 (red dashed line), G_A88 (black dashed line), G_B30 (blue solid line), G_B77 (red solid line), and G_B88 (black solid line). (*A*) Mean residue ellipticity (degrees per cm^2/dmol) is plotted vs. wavelength. Spectra were measured in 0.1 M potassium phosphate buffer (pH 7.2) using a 1-cm cylindrical cuvette at 25°C with [protein] = 5 μM. (*B*) Millidegrees at 222 nm are plotted vs. temperature in the range from 25°C to 100°C. The temperature profile was recorded by using a 1-cm cylindrical cuvette with a protein concentration of 5 μM in 0.1 M potassium phosphate buffer (pH 7.2).

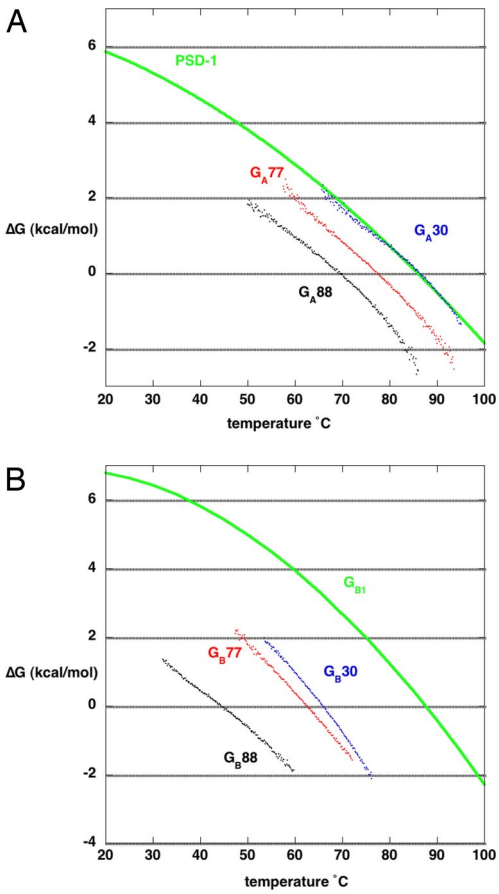


Fig. 4. Stability profiles of natural and mutant G_A and G_B proteins. (*A*) ΔG vs. temperature plots are shown for G_A30 (blue), G_A77 (red), and G_A88 (black). For reference, the stability curve for the parent protein PSD1 is shown in green. (*B*) ΔG vs. temperature plots are shown for G_B30 (blue), G_B77 (red), and G_B88 (black). For reference, the stability curve for the parent protein G_{B1} is shown in green.

populations). Restricting choices to category *i*, we mutated positions 1–8, 10, 15, 21, 22, 23, 26, 34, 44, 47, 53, and 54 in G_A30 to the corresponding amino acid in G_B30 and we mutated positions 11, 14, 16, 17, 36, 42, 46, and 48 in G_B30 to the corresponding amino acid in G_A30. The resulting heteromorphic pair (G_A77 and G_B77) was 77% identical (Fig. 2).

G_A77 and G_B77 were characterized by CD, thermal denaturation, ligand binding, and gel filtration as described for G_A30 and G_B30. These data are presented in Figs. 3 and 4. The stability of G_A77 has decreased by ≈1 kcal/mol relative to G_A30. The stability of G_B77 has decreased by ≈0.6 kcal/mol relative to G_B30. Both G_A77 and G_B77 are calculated to have a $\Delta G_{unfolding}$ of >4 kcal/mol at 25°C (Fig. 4). Both proteins show no decrease in binding affinity to their respective ligands relative to PSD1 and G_{B1} (by affinity chromatography) and are monomeric (by gel filtration). Despite the high degree of sequence identity (77%), the stability and binding characteristics of these two proteins remain similar to naturally occurring IgG and HSA binding domains. Both proteins are highly expressed in *Escherichia coli* in soluble form at 37°C.

In a final iteration, mutations at two sites from category *i* (51 and 52) were introduced into G_A77 and mutations at four sites from category *ii* were introduced into G_B77: G9L and L12A are compensating mutations and T18K and A20L are compensating mutations. The resulting proteins (G_A88 and G_B88) were 88% identical (Fig. 2). The CD and thermal denaturation data are presented in Figs. 3 and 4. The CD spectrum of G_A88 is essentially identical to PSD1, and the CD spectrum of G_B88 is essentially identical to that of native G_{B1}. The thermal denaturation midpoint is 70°C for G_A88 and 44°C for G_B88. The stability of G_A88 has decreased by ≈1 kcal/mol relative to G_A77. The stability of G_B88 has decreased by ≈2.4 kcal/mol relative to G_B77. G_A88 is calculated to have a $\Delta G_{unfolding}$ of ≈4 kcal/mol, and G_B88 is calculated to have a $\Delta G_{unfolding}$ of ≈2.0 kcal/mol at 25°C (Fig. 4). Both mutants unfold cooperatively. The small changes observed in the curvature of the stability curves for G_B77 and G_B88 may reflect small changes in $\Delta Cp_{unfolding}$.

HSA and IgG Sepharose columns were used to show that G_A88 binds only to HSA and G_B88 binds only to IgG. G_B88 eluted as a sharp peak in the flow-through fractions of the HSA column, and G_A88 eluted as a sharp peak in the flow-through fractions of the IgG column. G_B88 was completely retained through loading and washing steps on the IgG column, but G_A88 eluted from the HSA column in a broad peak after ≈10-column volumes. This elution profile was independent of flow rate indicating that binding is in rapid equilibrium. If the concentration of HSA on the column (0.17 mM) is taken as the free ligand concentration, then the K_D of G_A88 is estimated to be 20 μM. Although G_A88 binding affinity toward HSA has diminished relative to G_A77, it remains in the range observed for some naturally occurring G_A domains (7). Gel filtration over G25 and G75 columns showed that both proteins remain monomeric.

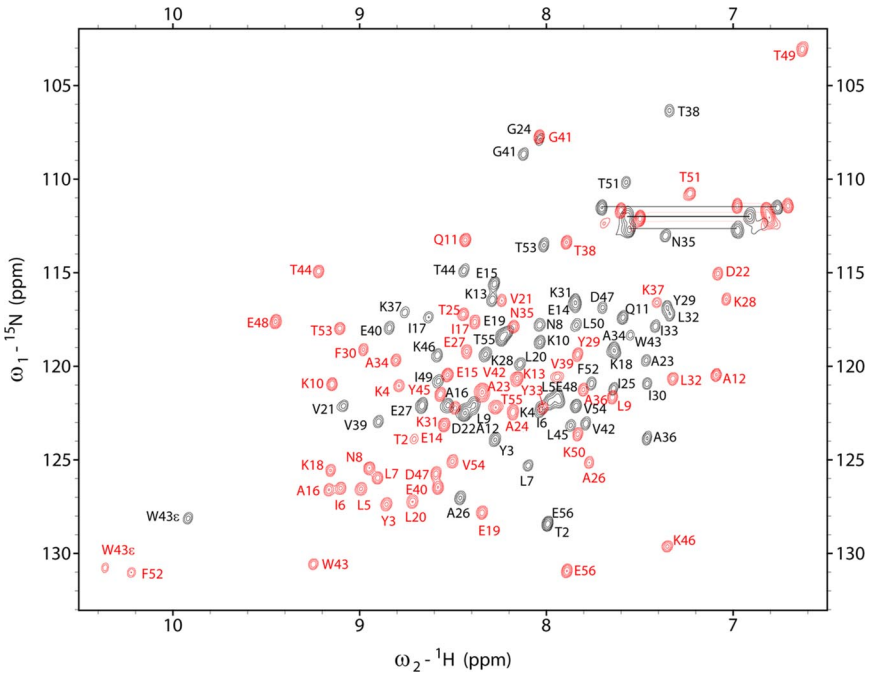


Fig. 5. ^{1}H, ^{15}N HSQC spectra. Main chain amide assignments are shown for G_A88 (black) and G_B88 (red). Forty-nine of the 56 aa in these two proteins are identical but have different chemical environments reflecting the two different folds. Amide proton signals for side chains are connected by the horizontal lines. Chemical shift index data for G_A88 and G_B88 are included in SI Tables 1 and 2.

Analysis of Folds by NMR. Previous studies have shown that native-like CD spectra and cooperative thermal denaturation profiles, although indicative of well defined tertiary structure, are not conclusive (20). Thus, to characterize their tertiary folds, G_A77, G_B77, G_A88, and G_B88 were isotopically labeled with ^{15}N and ^{13}C. ^{15}N, ^{1}H HSQC spectra of all four proteins are well dispersed with sharp line widths, indicative of folded proteins. Sequential assignments have been completed by using standard triple-resonance methods. The assigned spectra of G_A88 and G_B88 are shown in Fig. 5. The different chemical environments in the two folds result in little overlap of amide proton resonances in the two proteins, even though 49 of 56 aa are identical. Consensus chemical shift index comparisons [supporting information (SI) Tables 1 and 2] and analysis of diagnostic, long-range NOEs showed that G_A88 is very similar in fold to PSD1 and that G_B88 is very similar in fold to G_{B1}. High-resolution structures of G_A88 and G_B88 are nearing completion and will be published elsewhere.

Discussion

Previous studies have demonstrated that conformational preferences of many amino acid sequences are marginal. The structure of a short fragment within a protein can be completely context-dependent (29). Furthermore, monomeric proteins sometimes can assume alternative conformations in a multimeric form. This is observed in prion proteins and with monomeric and dimeric forms of the chemokine lymphotactin (30) and also with proteins designed to switch conformation and quaternary structure in the presence of a transition metal (31, 32). It has also been shown that five mutations to the core of G_B result in a conformational change to a stable, intertwined tetrameric state (33). It is clear from these examples that the energy of multimer formation or metal binding can drive conformational change.

We show here that the fold and function of a monomeric protein can depend on only 7 of 56 positions. The final identity of G_A88 and G_B88 results from nine initial identities, 17 identities created by mutations in G_B, 16 identities created by mutations in G_A (structured region only), and seven identities created by extension of the N terminus in G_A, for a total of 49 identities.

Thus, ≈30% of the structured amino acids in each member of the initial heteromorphic pair were mutated to produce the final 88% identity. Seven unique amino acids are preserved in each fold, and hence seven mutations are sufficient to shift the equilibrium from 99.9% 3-α fold to 97% α/β. The other 49 positions influence the 3-α to U and the α/β to U equilibria, but they can be tolerated in both folds and therefore provide a relatively neutral sequence background in which to observe an underlying fold-specific folding code. Clearly the complementary set of mutations in each member of the pair have to be very carefully selected to achieve this level of identity, but previous studies indicate that this result is probably not unique to this pair of proteins (17–21).

This fact has profound implications for understanding the protein folding code, as well as understanding how new folds and functions evolve. Protein stability is typically analyzed as a two-state reaction between a single folded state and a population of disordered, unfolded states. Our results suggest that the ΔG of alternative folded states may be much more favorable than generally recognized. This conclusion might explain a major difficulty in predicting the native fold by computational methods.

Our observation that a latent binding epitope can be engineered into a small protein without interfering with its native function suggests that the "folding problem" might be viewed from a new perspective. The vast array of conformations available to a polypeptide chain makes the folding problem difficult, but it also means that many latent functions could exist in this vast population of conformers. Because of the prevailing view that all folds except the native are highly improbable, the acquisition of new function by conformational switching also is judged to be improbable. Because the conformational preference for a 3-α vs. an α/β fold can depend on a delicate balance of critical interactions within a protein, however, a few mutations can result in a new conformation and the unmasking of new functionality.

The binary sequence space separating G_A88 and G_B88 comprises only 128 (2^7) different variations. In this population, most variants

will be predominantly unfolded. Some of these unfolded forms will likely have significant affinity for one and possibly even both ligands, if a sequence has significant propensity for both folds. Numerous examples exist of natural proteins that are largely unfolded unless bound to a ligand. We would suggest that transitional forms in the natural evolution of new folds may be predominantly unfolded (34). More than 30% of the eukaryotic proteome is predicted to be "natively unfolded." Innovation (unmasking of new function) could result from a few mutations that lower the propensity for the native fold but increase the propensity for an alternative one. Thus, the three characteristics of proteins that make the folding problem difficult (large conformational space, degenerate folding code, and small ΔG) may enable facile evolution of new folds and functions.

Methods

Protein Expression and Purification. To facilitate their rapid purification, G_A and G_B variants were cloned into the vector pG58, which encodes an engineered subtilisin prosequence as the N terminus of the fusion protein, and were purified by using an affinity-cleavage tag system that we developed (35). The system enabled the rapid, standardized purification of mutant proteins, even of low stability. Minimal medium (10) was used for ^{15}N and ^{13}C labeling. Soluble cell extract of prodomain fusion protein was injected on a 5-ml S189 column at 5 ml/min to allow binding and then washed with 10-column volumes of 100 mM KPO_4 (pH 7.2) to remove impurities (35). To cleave and elute the purified target protein, 6 ml of 100 mM KF/100 mM KPO_4 (pH 7.2) was injected at 0.5 ml/min. The purified protein was then dialyzed into 2 mM ammonium bicarbonate buffer (pH 7.0) and lyophilized.

CD. Lyophilized powder was resuspended in 100 mM KPO_4 (pH 7.2) for analysis by CD. Under these conditions all proteins are monomeric at concentrations up to 5 mg/ml (0.8 mM). This was demonstrated by gel filtration on G25 and G75 Sephadex. CD measurements were performed with a spectropolarimeter (model J-720; Jasco, Easton, MD) using water-jacketed quartz cells with path lengths of 1 cm on protein concentrations of 5 μM. The ellipticity results were expressed as mean residue ellipticity, $[\theta]$, degrees per $cm^2/dmol^{-1}$. Temperature-induced unfolding was performed in the temperature range between 25°C and 100°C in 1-cm cuvettes. Ellipticities at 222 nm were continuously monitored at a scanning rate of 1° per min. The fraction native is determined by subtracting unfolded baseline from the experimental CD signal and then dividing by the total CD difference between 100% folded and 0% folded at that temperature. Reversibility of the denaturations was confirmed by comparing the CD spectra at 25°C before melting and after heating to 100°C and cooling to 25°C. The temperature unfolding profiles measured by far-UV CD for G_A and G_B were converted to an apparent $\Delta G_{\text{unfolding}}$ and fit to a theoretical curve calculated by using the Gibbs–Helmholtz equation: $\Delta G_{\text{unfolding}} = \Delta H_o - T\Delta S_o + \Delta Cp(T - T_o - T\ln T/T_o)$, where $T_o = 298$ K and $\Delta Cp = 0.83$ kcal/°mol for G_B and 0.26 kcal/°mol for G_A (10, 36).

Binding to IgG and Human Serum Albumin. IgG and HSA were immobilized on GE HT1 columns containing NHS-activated agarose resin according to the manufacturer's instructions. Binding of PSD1, G_{B1}, and their mutants was carried out in 0.1 M KPO_4 (pH 7.2) by injecting 1 ml of a 1 mg/ml solution at 0.5 ml/min. Washing was with 15 ml of 0.1 M KPO_4 (pH 7.2) at 1 ml/min. Elution was with 6 ml of 0.5 M NaOAc (pH 3.0). All proteins except G_A88 were completely retained in the binding and washing steps.

NMR Spectroscopy. Lyophilized protein samples were dissolved in NMR buffer (50 mM NaPi/50 mM NaCl, pH 7.0) containing 10% D_2O. The final protein concentrations were in the range of 0.4–0.6 mM. NMR spectra were acquired on a AVANCE 600-MHz spectrometer (Bruker, Billerica, MA) equipped with a z axis gradient triple resonance ($^1H/^{13}C/^{15}N$) cryoprobe. Backbone resonance assignments were obtained from the following three-dimensional triple-resonance experiments recorded on $^{13}C/^{15}N$-labeled samples: HNCACB, CBCA(CO)NH, HBHA(CO)NH, and HNCO. All experiments were collected at 298 K. NMR spectra were processed by using nmrPipe (37) and analyzed with Sparky (38). Chemical shift index analysis was carried out with C_α, C_β, H_α, and CO assignments (39).

We thank David Rozak, John Moult, and Edward Eisenstein for their assistance and advice. This work was supported by National Institutes of Health Grant GM62154 (to J.O.).

1. Fahnestock SR, Alexander P, Nagle J, Filpula D (1986) *J Bacteriol* 167:870–880.
2. Falkenberg C, Bjork L, Åkerstrom B (1992) 31:1451–1457.
3. Frick IM, Wikstrom M, Forsen S, Drakenberg T, Gomi H, Sjobring U, Bjorck L (1992) *Proc Natl Acad Sci USA* 89:8532–8536.
4. Myhre EB, Kronvall G (1977) *Infect Immun* 17:475–482.
5. Reis KJ, Ayoub EM, Boyle MDP (1984) *J Immunol* 132:3098–3102.
6. de Chateau M, Holst E, Bjorck L (1996) *J Biol Chem* 271:26609–26615.
7. Rozak DA, Alexander PA, He Y, Chen Y, Orban J, Bryan PN (2006) *Biochemistry* 45:3263–3271.
8. He Y, Rozak DA, Sari N, Chen Y, Bryan P, Orban J (2006) *Biochemistry* 45:10102–10109.
9. Orban J, Alexander P, Bryan P (1994) *Biochemistry* 33:5702–5710.
10. Alexander P, Fahnestock S, Lee T, Orban J, Bryan P (1992) *Biochemistry* 31:3597–3603.
11. Gallagher TD, Alexander P, Bryan P, Gilliland G (1994) *Biochemistry* 33:4721–4729.
12. Gronenborn AM, Filpula DR, Essig NZ, Achari A, Whitlow M, Wingfield PT, Clore GM (1991) *Science* 253:657–661.
13. Achari A, Hale S, Howard AJ, Clore GM, Gronenborn AM, Hardman KD, Whitlow M (1992) *Biochemistry* 31:10449–10457.
14. Lian L-Y, Derrick JP, Sutcliffe MJ, Yang JC, Roberts GCK (1992) *J Mol Biol* 228:1219–1234.
15. Derrick JP, Wigley DB (1994) *J Mol Biol* 243:906–918.
16. Lattman EE, Rose GD (1993) *Proc Natl Acad Sci USA* 90:439–441.
17. Jones DT, Moody CM, Uppenbrink J, Viles JH, Doyle PM, Harris CJ, Pearl LH, Sadler PJ, Thornton JM (1996) *Proteins* 24:502–513.
18. Dalal S, Balasubramanian S, Regan L (1997) *Folding Des* 2:R71–R79.
19. Yuan SM, Clarke ND (1998) *Proteins* 30:136–143.
20. Blanco FJ, Angrand I, Serrano L (1999) *J Mol Biol* 285:741–753.
21. Dalal S, Regan L (2000) *Protein Sci* 9:1651–1659.
22. Lejon S, Frick IM, Bjorck L, Wikstrom M, Svensson S (2004) *J Biol Chem* 279:42924–42928.
23. Sauer-Eriksson AE, Keywegt GJ, Uhlen M, Jones TA (1995) *Structure (London)* 3:265–278.
24. Sloan DJ, Hellinga HW (1999) *Protein Sci* 8:1643–1648.
25. Livingstone JR, Spolar RS, Record TM (1991) *Biochemistry* 30:4237–4244.
26. Dalal S, Balasubramanian S, Regan L (1997) *Nat Struct Biol* 4:548–552.
27. Alexander PA, Rozak DA, Orban J, Bryan PN (2005) *Biochemistry* 44:14045–14054.
28. He Y, Yeh DC, Alexander P, Bryan PN, Orban J (2005) *Biochemistry* 44:14055–14061.
29. Minor DL, Jr, Kim PS (1996) *Nature* 380:730–734.
30. Kuloglu ES, McCaslin DR, Markley JL, Volkman BF (2002) *J Biol Chem* 277:17863–17870.
31. Ambroggio XI, Kuhlman B (2006) *J Am Chem Soc* 128:1154–1161.
32. Cerasoli E, Sharpe BK, Woolfson DN (2005) *J Am Chem Soc* 127:15008–15009.
33. Kirsten Frank M, Dyda F, Dobrodumov A, Gronenborn AM (2002) *Nat Struct Biol* 9:877–885.
34. Ward JJ, Sodhi JS, McGuffin LJ, Buxton BF, Jones DT (2004) *J Mol Biol* 337:635–645.
35. Ruan B, Fisher KE, Alexander PA, Doroshko V, Bryan PN (2004) *Biochemistry* 43:14539–14546.
36. Rozak DA, Orban J, Bryan PN (2005) *Biochim Biophys Acta* 1753:226–233.
37. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A (1995) *J Biomol NMR* 6:277–293.
38. Goddard TD, Kneller DG (2001) SPARKY (Univ of California, San Francisco), version 3.
39. Wishart DS, Sykes BD (1994) *J Biomol NMR* 4:171–180.
40. Derrick JP, Wigley DB (1992) *Nature* 359:752–754.